Saarland University

Faculty of Natural Sciences and Technology I

Department of Computer Science

Master thesis

Shattering Extremal Systems

submitted by
Shay Moran

submitted
03.09.12

Supervisor
Prof. Dr. Ami Litman

Supervisor
Prof. Dr. Kurt Mehlhorn

Reviewers
Prof. Dr. Kurt Mehlhorn
Prof. Dr. Ami Litman

Contents

1 Introduction

The *Shatters* relation and the *VC dimension* have been investigated since the early seventies (starting with [17, 18, 19]). These concepts have found numerous applications in statistics, combinatorics, learning theory and computational geometry.

Shattering extremal systems are set-systems with a very rich structure and many different characterizations. The goal of this paper is to elaborate on the structure of these systems. They were discovered several times and independently by several groups of researchers. Lawrence in [12] is the first who introduced them in his study of convex sets. Intrestingly, the definition he gave does not require the concept of shatters. Independently, Bollobás et al in [3] have discovered these systems, using the *shatters* relation (a.k.a *traces*); furthermore, Bollobás et al are the first who introduced the relation of *strongly-shatters* (a.k.a *strongly-traces*) and characterized Shattering extremal systems with the shatters and strongly-shatters relations. Dress et al, independently of Bollobas et al, have discovered the same characterization and established the equivalence to the characterization given by Lawrence.

It is an interesting phenomena that, in the past thirty years, different characterizations of these systems were discovered in different fields of pure and applied mathematics. Strangely, there seem to be no connections between the different groups that studied these systems. We hope to link together these different groups by this work and as a result to enhance the research of these systems. Here is a list of some of the contexts in which these systems were discovered: Functional analysis [13], Discrete-geometry ([12]), Phylogenetic Combinatorics ([5, 2]) and Extremal Combinatorics ([4, 3]). Moreover, as this work shows, the class of Shattering extremal systems naturally extends the class of *Maximum systems*[1]. Maximum systems occur in Geometry [8] and in Learning theory [6, 11, 16, 14, 15]. Thus, it is not unlikely to find usage of Shattering extremal systems in these fields.

1.1 Our results

We present new definitions of *shatters* and *strongly-shatters* that differ from each other only in the order of the quantifiers. We demonstrate that at least some of the known duality between these two concepts is due to this transpose of quantifiers. We shed additional light on this mysterious phenomena of duality via a mechanical transformation that, sometimes, translates claims to dual claims and proofs to dual proofs.

Two unary operators, which we denote ⊞ and ⊟, were extensively used under many different notations in the literature. These operators preserve the property of shattering extremality and the property of being a maximum system. Hence, these operarators were studied in both contexts. Many previous works observed a certain duality between these two operators. We demonstrate that this duality stems from the famous duality of Boolean algebra. Moreover, we reveal that a previously known lemma concerning duality is just the De-Morgan laws.

We present the famous *down shifting* operators in the terminology of oblivious sorting algorithms. The most common use of these operators is to apply a sequence

---

[1]set-systems that meet Sauer's inequality with equality

of them on a system until a fixed point is reached. Our presentation of down-shifting sheds light on this process and relates it to a basic problem in the theory of sorting.

We introduce new characterizations of the *shatters* and *strongly shatters* relations. One characterization is by the operators ⊣ and ⊢ and the other characterization is by down-shifting operators. As a corollary we obtain simple proofs to two known characterizations of Shattering extremal systems via commutativity between ⊣ and ⊢ and commutativity between down-shifting operators. Moreover, we introduce a weaker form of Shattering-Extremality and establish equivalences of this form with a weaker form of the above commutativities.

We show that, similarly to Maximum systems, Shattering extremal systems occur naturally in Geometry as a combinatorial feature of an arrangement of oriented hyperplanes. We give a geometrical interpretation of *shattering* and *strong-shattering* in this setup. We also introduce the class of *"Convex systems"* which form a simple and natural generalization of the classes studied by Welzl ([8]) and of the classes studied by Lawrence ([12]). We pose as an open question whether every Shattering Extremal system is a convex system..

Finally, we use this theory to develop a machinery for proving certain equalities and inequalities concerning undirected graphs and their orientations. As an example of this machinery we will prove the following inequality:
Let $G = (V, E)$ be an undirected graph and let s, t be two distinct vertices in V. Then, the number of orientations of E that contain a directed path from s to t is the same as the number of subgraphs[2] of G that contain an undirected $s - t$ path.
One advantage of this machinery is that it demonstrates an application of *Strong shattering* which is exceptional since most applications of this theory use *shattering*.

---

[2]i.e. subsets of E

2 Shattering Extremal systems - A first date

2.1 Basic definitions

This work studies finite set-systems and thus all structures discussed in this paper are finite. A standard definition of a set-system is a pair (T,X) where X is a set and $T \subseteq \mathcal{P}(X)$. For reasons that will become clear later, we choose a different definition for set-systems. Also, we prefer to call these structures "systems" rather than "set-systems".

Definition 2.1 (System). *A **system** is a pair, $\mathbb{S} = \langle S, \{0,1\}^X \rangle$, where X is a set and $S \subseteq \{0,1\}^X$. For a system $\mathbb{S} = \langle S, \{0,1\}^X \rangle$:*

- *Let $S(\mathbb{S})$, $\dim(\mathbb{S})$ and $C(\mathbb{S})$ denote S, X and $\{0,1\}^X$ respectively.*

- *Let $|\mathbb{S}|$ denote $|S|$.*

- *The **complement** of $\mathbb{S}$, denoted $\neg\mathbb{S}$, is the system defined by:*

$$\neg\mathbb{S} \stackrel{def}{=} \langle C(\mathbb{S}) - S(\mathbb{S}), C(\mathbb{S}) \rangle.$$

- *The elements of $C(\mathbb{S})$ are called vertices*

Note that for every set X there are exactly $2^{2^{|X|}}$ systems, $\mathbb{S}$, with $C(\mathbb{S}) = \{0,1\}^X$. In particular, there are exactly two systems for $X = \emptyset$, one of them is empty and the other contains one vertex. Let $\mathcal{K}_0$ denote the former and let $\mathcal{K}_1$ denote the latter system.

Definition 2.2. *Let f, g be arbitrary functions and let $A \subseteq Dom(f) \cap Dom(g)$.*

- *f **agrees with** g **on** A means that $f(x) = g(x)$ for all $x \in A$.*

- *f **agrees with** g means that f agrees with g on $Dom(f) \cap Dom(g)$*

Definition 2.3 (Cube). *Let X be a set and let $Y \subseteq X$. A Y**-cube** of $\{0,1\}^X$ is an equivalence class of the following equivalence relation on $\{0,1\}^X$: "u agrees with v on $X - Y$"*

- *Let C be a Y-cube of $\{0,1\}^X$. The **dimensions of** C are defined to be the set Y and denoted by $\dim(C)$.*

Note that there are $2^{|X-Y|}$ Y-cubes, they are disjoint and cover all of $\{0,1\}^X$.

2.2 Lopsided systems

Lawrence introduced the concept of lopsided systems as follows

Definition 2.4 (Lawrence [12]). *A system $\mathbb{S}$ is called **lopsided** if for every $\{X', X''\}$, a partitioning of $\dim(\mathbb{S})$, exactly one of the following statements is true:*

1. *$S(\mathbb{S})$ contains an X'-cube*

2. *$S(\neg\mathbb{S})$ contains a X''-cube*

The following important property of lopsided set-sytems is an immediate corollary of the definition:

Theorem 2.5 (Complementing). *A system, $\mathbb{S}$, is lopsided if and only if $\neg\mathbb{S}$ is lopsided.*

2.3 Shatters and Strongly-Shatters

Definition 2.6 (Merging)**.** *Let $f \in \{0,1\}^X$ and $g \in \{0,1\}^Y$ where X,Y are disjoint. We define $f \star g$ to be the unique function in $\{0,1\}^{X \cup Y}$ that agrees with both f and g.*

Note that $\star$ is a commutative and associative operation.
Let $\mathbb{S}$ be a system, let $X = \dim(\mathbb{S})$ and let $Y \subseteq \dim(\mathbb{S})$.

Definition 2.7 (Shatters)**.** *We say that $\mathbb{S}$ **shatters** Y if*

$$(\forall f \in \{0,1\}^Y)(\exists g \in \{0,1\}^{X-Y}) : g \star f \in S(\mathbb{S})$$

Definition 2.8 (Strong-shatters)**.** *We say that $\mathbb{S}$ **strongly shatters** Y if*

$$(\exists g \in \{0,1\}^{X-Y})(\forall f \in \{0,1\}^Y) : g \star f \in S$$

That is, our definitions of *strongly shatters* and *shatters* are identical except of the order of the two quantifiers. These definitions highlight the duality that was observed by previous works. For example, a straightforward application of predicate calculus on the above definitions gives the next lemma:

Lemma 2.9 ([3, 1])**.** *Let $\mathbb{S}$ be a system and let $\{X',X''\}$, be a partitioning of $\dim(\mathbb{S})$. Then exactly one of the following statements is true:*

1. *$\mathbb{S}$ shatters X'*

2. *$\neg \mathbb{S}$ strongly shatters X''*

Note the similarity between Lemma 2.9 and definition 2.4. That is, the former is derived from the latter by replacing "*shatters*" in Statement 1 with "*strongly-shatters*".
A natural variant of "*Lopsided*" is the following concept:

Definition 2.10 (Dual lopsided system)**.** *A system $\mathbb{S}$ is called **Dual lopsided** if for every $\{X',X''\}$, a partitioning of $\dim(\mathbb{S})$, exactly one of the following statements is true:*

1. *$\mathbb{S}$ shatters X'*

2. *$\neg \mathbb{S}$ shatters X''*

Two subsets of $\mathcal{P}(\dim(\mathbb{S}))$ are associated with $\mathbb{S}$:

- $\mathrm{str}(\mathbb{S}) \stackrel{def}{=} \{Y \subseteq \dim(\mathbb{S}) \mid Y \text{ is shattered by } \mathbb{S}\}$

- $\mathrm{sstr}(\mathbb{S}) \stackrel{def}{=} \{Y \subseteq \dim(\mathbb{S}) \mid Y \text{ is strongly shattered by } \mathbb{S}\}$

Obviously, both $\mathrm{str}(\mathbb{S})$ and $\mathrm{sstr}(\mathbb{S})$ are closed under subset relation and $\mathrm{sstr}(\mathbb{S}) \subseteq \mathrm{str}(\mathbb{S})$. This work studies systems for which $\mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S})$.

2.4 Shattering extremal systems

Definition 2.11 (Shattering-Extremal system)**.** $\mathbb{S}$ *is* **Shattering-Extremal**[3] *(in abbreviation: SE) if it satisfies*

$$\mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S}).$$

The property "$\mathbb{S}$ is SE" is sometimes refered to by the expression "$SE(\mathbb{S})$".

2.5 Shattering extremality = lopsidedness

This section proves the following theorem:

Theorem 2.12. *The following statements are equivalent for a system $\mathbb{S}$:*

1. *$SE(\mathbb{S})$*

2. *$SE(\neg\mathbb{S})$*

3. *$\mathbb{S}$ is lopsided*

4. *$\mathbb{S}$ is dual lopsided*

The equivalence of the first three statements was proven by Dress et al in [5].

Lemma 2.13. *Let $\mathbb{S}$ be a system. $\mathbb{S}$ is lopsided $\iff$ $\mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S})$*

Proof.

$$
\begin{aligned}
Lopsided(\mathbb{S}) &\iff \forall X \in \mathcal{P}(\dim(\mathbb{S})) : \dim(\mathbb{S}) - X \notin \mathrm{sstr}(\neg\mathbb{S}) \leftrightarrow X \in \mathrm{sstr}(\mathbb{S}) && \text{by definition} \\
&\iff \forall X \in \mathcal{P}(\dim(\mathbb{S})) : X \in \mathrm{str}(\mathbb{S}) \leftrightarrow X \in \mathrm{sstr}(\mathbb{S}) && \text{by Lemma 2.9} \\
&\iff \mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S}).
\end{aligned}
$$

$\square$

As noted earlier, there is a certain duality in this theory between shattering and strong shattering. This duality is manifested by a certain mechanical tranformation on text written in "mathematical english". The texts of interest are lemmas and proofs and we refer to the transformed texts as dual lemmas and dual proofs. This dual transformation swaps few words and symbols as follows: It swaps the pair "str" and "sstr", the pair "Lopsided" and "Dual lopsided", the pair "$\subseteq$" and "$\supseteq$" and the pair "$\leq$" and "$\geq$". This dual transformation is useful since, sometimes, the dual of a true lemma is true. Moreover, sometimes the dual of a proof is a valid proof of the dual lemma. Lemma 2.13 and its proof are an example to this phenomenon.

Lemma 2.14 (Dual of Lemma 2.13)**.** *Let $\mathbb{S}$ be a system. Then*

$$\mathbb{S} \text{ is dual lopsided} \iff \mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S}).$$

It is easy to verify that the dual of the proof of Lemma 2.13 is a valid proof of Lemma 2.14.

These two lemmas establish equivalence between items $1, 3, 4$ in Theorem 2.12. The rest of the theorem is derived from Theorem 2.5.

---

[3]This is not the standard definition of Shattering Extremal systems. The standard definition and its equivalence to this definition will be discussed in Section 3.3

3 Inequalities

This chapter presents basic inequalities which involve the concepts of *Shattering* and *Strong-Shattering* and discusses their relation with *SE* systems.

3.1 VC-dimension and Sauer's lemma

Definition 3.1 (VC-dimension)**.** *The **VC-dimension** (Vapnik Chervonenkis dimension) of a system $\mathbb{S}$, denoted $\mathrm{vc}(\mathbb{S})$, is the cardinality of the largest[4] subset that is shattered by it.*

The well-known "Sauer lemma" was proved in the 70's by Sauer ([17]), Shelah ([18]) and Vapnik and Chervonenkis ([19]):

Theorem 3.2 (Sauer's lemma)**.** *Let $\mathbb{S}$ be a system with $|\dim(\mathbb{S})| = n$. Then*

$$|\mathbb{S}| \leq \sum_{i=0}^{\mathrm{vc}(\mathbb{S})} \binom{n}{i}$$

A proof for this lemma will be given in the next section.

3.2 The Sandwich theorem

The following generalization of Sauer's lemma is the result of an accumulated work by several authors. Different parts of this theorem were proven independently several times. (Pajor [13], Bollobás and Radcliffe [3], Dress [5] and Holzman and Aharoni [1, 9])

Theorem 3.3 (Sandwich theorem [13, 3, 5, 1, 9])**.** *For a system $\mathbb{S}$:*

$$|\mathrm{sstr}(\mathbb{S})| \leq |\mathbb{S}| \leq |\mathrm{str}(\mathbb{S})|$$

This section discusses and proves this theorem. The following trivial fact will be useful. For any $a', b', a'', b'' \in \mathbb{R}$:

$$a' \leq b' \text{ and } a'' \leq b'' \text{ and } a' + b' = a'' + b'' \implies a' = b' \text{ and } a'' = b'' \qquad (1)$$

---

[4]As a special case, $\mathrm{vc}(\mathbb{S}) = -1$ when $S(\mathbb{S}) = \emptyset$

Lemma 3.4. *Let* $\mathbb{S}, \mathbb{S}', \mathbb{S}''$ *be systems that satisfy the following:*

1. $|\mathbb{S}'| + |\mathbb{S}''| = |\mathbb{S}|$

2. $|\mathrm{str}(\mathbb{S}')| + |\mathrm{str}(\mathbb{S}'')| \le |\mathrm{str}(\mathbb{S})|$

3. $|\mathbb{S}'| \le |\mathrm{str}(\mathbb{S}')|$, $|\mathbb{S}''| \le |\mathrm{str}(\mathbb{S}'')|$

Then:

(a) $|\mathbb{S}| \le |\mathrm{str}(\mathbb{S})|$

(b) $|\mathbb{S}| = |\mathrm{str}(\mathbb{S})| \iff$ *Assumptions* 2 *and* 3 *in the premise of the lemma are satisfied with equality*

Remark: Assumption 3 is always true as will be proved later

Proof. Regarding Conclusion (a):

$$\begin{aligned} |\mathrm{str}(\mathbb{S})| &\ge |\mathrm{str}(\mathbb{S}')| + |\mathrm{str}(\mathbb{S}'')| && \text{by Assumption 2} \\ &\ge |\mathbb{S}'| + |\mathbb{S}''| && \text{by Assumption 3} \\ &= |\mathbb{S}| && \text{by Assumption 1} \end{aligned}$$

.

Considering Conclusion (b): The right to left direction is trivial. Assume $|\mathbb{S}| = |\mathrm{str}(\mathbb{S})|$. Therefore, the above chain of inequalities holds when the symbol '$\ge$' is replaced with '$=$'. This implies equality in Assumption 2. Fact (1) implies equality in Assumption 3. $\qquad\square$

Definition 3.5. *Let* $\mathbb{S}$ *be a system, let* $x \in \dim(\mathbb{S})$. *The* **restrictions of** $\mathbb{S}$ **associated with** x *are the following two systems:*

- $\langle \{ f \restriction_{\dim(\mathbb{S}) - \{x\}} \ : \ f \in \mathbb{S}, \ f(x) = 0 \}, \{0,1\}^{\dim(\mathbb{S}) - \{x\}} \rangle$

- $\langle \{ f \restriction_{\dim(\mathbb{S}) - \{x\}} \ : \ f \in \mathbb{S}, \ f(x) = 1 \}, \{0,1\}^{\dim(\mathbb{S}) - \{x\}} \rangle$

Let $\mathbb{S}$ be a system, let $x \in \dim(\mathbb{S})$ and let $\{\mathbb{S}', \mathbb{S}''\}$ be the pair of restrictions of $\mathbb{S}$ associated with x. We will prove that $|\mathbb{S}| \le |\mathrm{str}(\mathbb{S})|$ by showing that these particular $\mathbb{S}, \mathbb{S}', \mathbb{S}''$ satisfy the premise of Lemma 3.4.

Definition 3.6. *Let* A, B *be sets and* x *be an object. Let* $\oplus_x$ *denote the following operator:*

$$A \oplus_x B \overset{def}{=} A \cup B \cup \big\{ c \cup \{x\} \ : \ c \in A \cap B \big\}$$

The following two lemmas are straightforward.

Lemma 3.7. *Let* A, B, *be sets and let* x *be an object such that* $x \notin T$ *for every* $T \in A \cup B$. *Then*

$$|A \oplus_x B| = |A| + |B|$$

Lemma 3.8. *Let* $\mathbb{S}$ *be a system, let* $x \in \dim(\mathbb{S})$ *and let* $\{\mathbb{S}', \mathbb{S}''\}$ *be the pair of restrictions of* $\mathbb{S}$ *associated with x. Then:*

$$\text{str}(\mathbb{S}') \oplus_x \text{str}(\mathbb{S}'') \subseteq \text{str}(\mathbb{S})$$

Lemma 3.9. *Let* $\mathbb{S}$ *be a system. Then:*

$$|\mathbb{S}| \leq |\text{str}(\mathbb{S})|.$$

Proof. We prove the claim by induction on $\dim(\mathbb{S})$. The case of $\dim(\mathbb{S}) = \emptyset$ is trivial. Otherwise, pick $x \in \dim(\mathbb{S})$ and let $\{\mathbb{S}', \mathbb{S}''\}$ be the pair of restrictions of $\mathbb{S}$ associated with x. It is enough to show that $\mathbb{S}', \mathbb{S}''$ satisfy the three Assumptions of Lemma 3.4. Assumption 1 is immediate. Assumption 2 follows from Lemmas 3.7 and 3.8 Assumption 3 follows by the induction hypothesis. $\qquad\square$

To see that Lemma 3.9 generalize Sauer lemma, note that by the definition of vc dimension:

$$\text{str}(\mathbb{S}) \subseteq \{Y \subseteq \dim(\mathbb{S}) \ : \ |Y| \leq \text{vc}(\mathbb{S})\}.$$

Hence, Sauer lemma (Theorem 3.2) is an easy conclusion of Lemma 3.9.

To establish the Sandwich theorem it remains to show that $|\text{sstr}(\mathbb{S})| \leq |\mathbb{S}|$ and this is accomplished via applying the Duality tranformation on the proofs of Lemmas 3.4 and 3.8 and 3.9. It is easy to verify that the duals of our proofs for these lemmas are valid[5]. The dual of Lemma 3.9 gives the desired inequality, namely,

$$|\text{sstr}(\mathbb{S})| \leq |\mathbb{S}|.$$

This finishes the proof of Theorem 3.3.

Our proof for the Sandwich Theorem demonstrates the usefulness of the Duality Transformation. However, it is important to write some lines about the fragile nature of this tranformation. This transformation works only for some lemmas/theorems. It is not hard to find (true) claims whose dual claims are not true. Moreover, as we will see in the next section (Lemma 3.11), there are some cases in which the dual of a lemma is true but the dual of the presented proof is not valid. Thus, when using this transformation, one has to take special care and to check the validity of the arguments being transformed.

3.3   SE systems and the Sandwich theorem

The inequalities in the sandwich theorem suggest two more extremal properties of systems, namely:

- Systems $\mathbb{S}$ for which $|\text{sstr}(\mathbb{S})| = |\mathbb{S}|$

- Systems $\mathbb{S}$ for which $|\mathbb{S})| = |\text{str}(\mathbb{S})|$

---

[5]Note that the Dual transformation of a proof transforms only the text explicitly written. It does not translate implicit arguments

The following result concerning the equivalence of these extremal properties was proven independently in [2], and in [3]:

Theorem 3.10. *Let $\mathbb{S}$ be a system. The following statements are equivalent:*

1. $\mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S})$

2. $|\mathrm{sstr}(\mathbb{S})| = |\mathbb{S}|$

3. $|\mathbb{S}| = |\mathrm{str}(\mathbb{S})|$

This section discusses and proves Theorem 3.10.

Lemma 3.11. *For every system $\mathbb{S}$:* $|\mathrm{str}(\mathbb{S})| = |\mathbb{S}| \implies \mathrm{sstr}(\mathbb{S}) = \mathrm{str}(\mathbb{S})$

Proof. The lemma is proven by induction on $|\mathrm{dim}(\mathbb{S})|$. Let $B \in \mathrm{str}(\mathbb{S})$. It is enough to show that $B \in \mathrm{sstr}(\mathbb{S})$. The case of $B = \mathrm{dim}(\mathbb{S})$ is easy. Otherwise, pick $x \notin B$ and let $\mathbb{S}', \mathbb{S}''$ be the two restrictions of $\mathbb{S}$ associated with x. By Lemma 3.4, $\mathbb{S}', \mathbb{S}''$ satisfy the premise of the current lemma. By the induction hypothesis:

$$\mathrm{sstr}(\mathbb{S}') = \mathrm{str}(\mathbb{S}') \tag{2}$$

$$\mathrm{sstr}(\mathbb{S}'') = \mathrm{str}(\mathbb{S}'') \tag{3}$$

Moreover, Lemma 3.4 also implies that $|\mathrm{str}(\mathbb{S}')| + |\mathrm{str}(\mathbb{S}'')| = |\mathrm{str}(\mathbb{S})|$. This, combined with Lemmas 3.7 and 3.8, implies that

$$\mathrm{str}(\mathbb{S}') \oplus_x \mathrm{str}(\mathbb{S}'') = \mathrm{str}(\mathbb{S}) \tag{4}$$

Therefore:

$$
\begin{aligned}
B \in \mathrm{str}(\mathbb{S}) &\implies B \in \mathrm{str}(\mathbb{S}') \oplus_x \mathrm{str}(\mathbb{S}'') && \text{by Equation 4}\\
&\implies B \in \mathrm{str}(\mathbb{S}') \cup \mathrm{str}(\mathbb{S}'') && \text{since } x \notin B\\
&\implies B \in \mathrm{sstr}(\mathbb{S}') \cup \mathrm{sstr}(\mathbb{S}'') && \text{by Equations (2) and (3)}\\
&\implies B \in \mathrm{sstr}(\mathbb{S}).
\end{aligned}
$$

$\square$

The dual of Lemma 3.11 is true. However, we dont know a proof for Lemma 3.11 whose dual is valid[6]. Therefore, we present a different proof for the dual lemma.

Lemma 2.9 implies that for all $\mathbb{S}$: $|\mathrm{str}(\mathbb{S})| + |\mathrm{sstr}(\neg\mathbb{S})| = 2^{|\mathrm{dim}(\mathbb{S})|}$. Clearly, $|\mathbb{S}| + |\neg\mathbb{S}| = 2^{|\mathrm{dim}(\mathbb{S})|}$. This implies:

Lemma 3.12. *For every system $\mathbb{S}$, the following statements are equivalent:*

1. $|\mathbb{S}| = |\mathrm{str}(\mathbb{S})|$

2. $|\mathrm{sstr}(\neg\mathbb{S})| = |\neg\mathbb{S}|$

---

[6]The dual of our proof of Lemma 3.11 is not valid as observed, for example, by the dual of the second line in the proof: "Let $B \in \mathrm{sstr}(\mathbb{S})$. It is enough to show that $B \in \mathrm{str}(\mathbb{S})$... ". This argument is obviously wrong.

Lemma 3.13 (Dual of Lemma 3.11). *For every system* $\mathbb{S}$:

$$|\text{sstr}(\mathbb{S})| = |\mathbb{S}| \implies \text{sstr}(\mathbb{S}) = \text{str}(\mathbb{S}).$$

Proof.

$$
\begin{aligned}
|\text{sstr}(\mathbb{S})| = |\mathbb{S}| &\implies |\text{str}(\neg\mathbb{S})| = |\neg\mathbb{S}| && \text{by Lemma 3.12} \\
&\implies \text{sstr}(\neg\mathbb{S}) = \text{str}(\neg\mathbb{S}) && \text{by Lemma 3.11} \\
&\implies \text{sstr}(\mathbb{S}) = \text{str}(\mathbb{S}) && \text{by Theorem 2.12}
\end{aligned}
$$

$$\cdot$$

$\square$

The Sandwich Theorem and Lemmas 3.11, 3.13 finish the proof of Theorem 3.10.
By Theorem 3.10 it follows that any of the 3 properties mentioned there, may serve as a definition of *Shattering Extremality*.
The most common definition in the literature is $|\mathbb{S}| = |\text{str}(\mathbb{S})|$. (See [3, 1, 9])
We prefer $\text{sstr}(\mathbb{S}) = \text{str}(\mathbb{S})$ as this definition is invariant under duality.

3.4 A recursive characterization of *SE*

Theorem 3.14. *Let* $\mathbb{S}$ *be a system, let* $x \in \dim(\mathbb{S})$ *and let* $\{\mathbb{S}', \mathbb{S}''\}$ *be the pair of restrictions of* $\mathbb{S}$ *associated with x. Then the following statements are equivalent:*

1. *SE*$(\mathbb{S})$.

2. *SE*$(\mathbb{S}')$ *and SE*$(\mathbb{S}'')$ *and* $\text{str}(\mathbb{S}') \oplus_x \text{str}(\mathbb{S}'') = \text{str}(\mathbb{S})$.

3. *SE*$(\mathbb{S}')$ *and SE*$(\mathbb{S}'')$ *and* $\text{sstr}(\mathbb{S}') \oplus_x \text{sstr}(\mathbb{S}'') = \text{sstr}(\mathbb{S})$.

Proof. The proof is divided into two parts.
First part: 1 $\iff$ 2. By Lemmas 3.7 and 3.8 it follows that

$$\text{str}(\mathbb{S}') \oplus_x \text{str}(\mathbb{S}'') = \text{str}(\mathbb{S}) \iff |\text{str}(\mathbb{S}')| + |\text{str}(\mathbb{S}'')| = |\text{str}(\mathbb{S})|.$$

This, combined with Conclusion (b) of Lemma 3.4 and Theorem 3.10, concludes the first part of this proof.
Second part: 1 $\iff$ 3. This part is the dual of the first part. It is easy to verify that the dual of the proof given for the first part is valid. $\square$

3.5 Maximum systems

The class of Maximum systems is the "famous relative" of the class of SE-systems.

Definition 3.15 (Maximum systems). $\mathbb{S}$ *is called a* **Maximum system** *if it meets Sauer's lemma with equality:*

$$|\mathbb{S}| = \sum_{i=0}^{\text{vc}(\mathbb{S})} \binom{n}{i}$$

Maximum systems and their applications in machine learning have been discussed in [16, 14, 15, 11, 6]. The relation between maximum systems and arrangements of hyper-planes in a Euclidean space have been discussed in [8]

SE systems are an extension of maximum systems (i.e every maximum system is SE). Moreover, SE systems seem to possess all "nice properties" of maximum systems and some other "nice properties" that maximum systems do not have. (See for example Theorem 4.4)

4 Operations on systems

4.1 Pre-systems and restrictions

Let $\mathbb{S}$ be a system and let $C \subseteq C(\mathbb{S})$ be a cube. Consider the structure $\langle S(\mathbb{S}) \cap C, C \rangle$. According to our terminology this structure is not a system. We refer to it as a "pre-system". Namely:

Definition 4.1 (pre-system). *Let X be a set. A pre-system is a pair $\langle S, C' \rangle$ where C' is a cube of $\{0,1\}^X$ and $S \subseteq C'$*

Clearly, every pre-system $\langle S, C' \rangle$ can be translated to a system by restricting its vertices to $\dim(C')$ as follows:

Definition 4.2 (Normalizing). *Let X be a set and let $C \subseteq \{0,1\}^X$ be a cube. Define $n : C \to \{0,1\}^{\dim(C)}$ by*

$$n(f) = f \restriction_{\dim(C)}$$

This mapping is called the normalizing of C.

this mapping is naturally extended to subsets of C as follows, Let $S \subseteq C$. Then $n(S) \overset{def}{=} \{n(f) \mid f \in S\}$; and to pre-systems. Let $\mathbb{S} = \langle S, C \rangle$ be a pre-system. Then $n(\mathbb{S}) \overset{def}{=} \langle n(S), \{0,1\}^{\dim(C)} \rangle$. Note that the normalization of a pre-system is a system.

Definition 4.3 (restriction). *Let $\mathbb{S}$ be a system, let $C' \subseteq C(\mathbb{S})$ be a cube. Let $\mathbb{S} \restriction_{C'}$ denote the normalization of the pre-system $\langle S(\mathbb{S}) \cap C', C' \rangle$. This system is called the **restriction** of $\mathbb{S}$ to C'*

4.1.1 Restrictions preserve SE

The following important property of *SE* systems is an immediate corollary of Lemma 3.14

Theorem 4.4. *Let $\mathbb{S}$ be a system, and let C be a cube. Then:*

$$SE(\mathbb{S}) \implies SE(\mathbb{S} \restriction_C)$$

The property of "being a maximum system" is not preserved by restrictions. Therefore this theorem is an example of the advantage of SE systems over maximum systems.

4.2 Boolean operations

Definition 4.5 (Boolean operations). *Let $\mathbb{A}, \mathbb{B}$ be two systems with $\dim(\mathbb{A}) = \dim(\mathbb{B}) = X$. The **union** and **intersection** of $\mathbb{A}, \mathbb{B}$ is defined as follows:*

1. $\mathbb{A} \uplus \mathbb{B} \overset{def}{=} \langle S(\mathbb{A}) \cup S(\mathbb{B}), \{0,1\}^X \rangle$

2. $\mathbb{A} \Cap \mathbb{B} \overset{def}{=} \langle S(\mathbb{A}) \cap S(\mathbb{B}), \{0,1\}^X \rangle$

Note that if $\dim(\mathbb{A}) \neq \dim(\mathbb{B})$ then these operations are undefined and that both $\uplus$, $\Cap$ are associative and commutative binary operations. These two operations and the complement operator (2.1) satisfy the relations of the (corresponding) operations in a boolean algebra.

As an example, the following is a trivial variant of the famous De-Morgan's laws:

1. $\neg(\mathbb{A} \uplus \mathbb{B}) = (\neg\mathbb{A}) \Cap (\neg\mathbb{B})$

2. $\neg(\mathbb{A} \Cap \mathbb{B}) = (\neg\mathbb{A}) \uplus (\neg\mathbb{B})$

These operations are used as unary operators in the standard way: Let $\mathcal{T}$ be a collection of systems such that $\forall \mathbb{S}, \mathbb{S}' \in \mathcal{T} : \dim(\mathbb{S}) = \dim(\mathbb{S}')$. $\uplus \mathcal{T}$ denotes the union over all elements of $\mathcal{T}$ and $\Cap \mathcal{T}$ denotes the intersection over all elements of $\mathcal{T}$.

4.3   The Boolean operators $\Cap_Y$ and $\uplus_Y$

Definition 4.6 ($\Cap_Y$ and $\uplus_Y$)**.** *Let Y be a set. The following operators, $\Cap_Y, \uplus_Y$ are defined only on systems $\mathbb{S}$ such that $Y \subseteq \dim(\mathbb{S})$. In this case:*

1.
$$\underset{Y}{\Cap}(\mathbb{S}) \overset{def}{=} \Cap \{S{\restriction}_C \mid C \text{ is a cube with } \dim(C) = \dim(\mathbb{S}) - Y\}$$

2.
$$\underset{Y}{\uplus}(\mathbb{S}) \overset{def}{=} \uplus \{S{\restriction}_C \mid C \text{ is a cube with } \dim(C) = \dim(\mathbb{S}) - Y\}$$

For an operator $q \in \{\uplus_Y, \Cap_Y\}$, let $\dim(q) \overset{def}{=} Y$.

The operators $\Cap_Y, \uplus_Y$ are extensively used under many different notations in the literature. Most works noted the duality between these opersators. However, our notations highlights the fact that this duality stems from the famous duality of Boolean algebra. For example, several works ([2]- page 674,(11) and [8]- observation 22) observed the following lemma, Lemma 4.7, but didn't notice that this lemma is the famous De-morgan laws.

Lemma 4.7 (De-Morgan)**.** *Let Y be a set, then:*

- $\Cap_Y \circ \neg = \neg \circ \uplus_Y$

- $\uplus_Y \circ \neg = \neg \circ \Cap_Y$

4.3.1   Analogues of theorems and lemmas

Definition 4.8. *Let $\mathbb{S}$ be a system and let $x \in \dim(\mathbb{S})$. The **derivatives of $\mathbb{S}$ associated with** x, are the set $\{\uplus_{\{x\}}(\mathbb{S}), \Cap_{\{x\}}(\mathbb{S})\}$.*

Several lemmas and theorems from chapter 2 contain the phrase "let $\{\mathbb{S}', \mathbb{S}''\}$ be the pair of restrictions of $\mathbb{S}$ associated with x". Namely: Lemma 3.8 and Theorem 3.14. The only properties of the above $\mathbb{S}', \mathbb{S}''$ that were used in proving these lemmas and theorems are: $|\mathbb{S}'| + |\mathbb{S}''| = |\mathbb{S}|$ and the fact that $\{\mathbb{S}', \mathbb{S}''\}$ satisfy Lemma 3.8. It is easy to verify that the pair of derivatives (associated with x) of a system $\mathbb{S}$ also have these two properties. Hence these theorems and lemmas hold when the above phrase is replaced with the phrase "let $\{\mathbb{D}', \mathbb{D}''\}$ be the two derivatives of $\mathbb{S}$ associated with x". Hence, the Sandwich Theorem may be proved by induction via the derivatives rather than the restrictions. An important result that is established this way is the analogue of Theorem 3.14:

Theorem 4.9. *Let $\mathbb{S}$ be a system, let $x \in \dim(\mathbb{S})$ and let $\mathbb{D}', \mathbb{D}''$ be the pair of derivatives of $\mathbb{S}$ associated with x. Then the following statements are equivalent:*

1. *$SE(\mathbb{S})$.*

2. *$SE(\mathbb{D}')$ and $SE(\mathbb{D}'')$ and $\mathrm{str}(\mathbb{D}') \oplus_x \mathrm{str}(\mathbb{D}'') = \mathrm{str}(\mathbb{S})$.*

3. *$SE(\mathbb{D}')$ and $SE(\mathbb{D}'')$ and $\mathrm{sstr}(\mathbb{D}') \oplus_x \mathrm{sstr}(\mathbb{D}'') = \mathrm{sstr}(\mathbb{S})$.*

It is quite surprising that these claims concerning $\mathbb{D}', \mathbb{D}''$ are established using arguments that are symmetric in $\mathbb{D}'$ and $\mathbb{D}''$.

4.3.2 $\boxminus_Y$ and $\boxplus_Y$ preserve *SE*

The following theorem was proven by several authors:

Theorem 4.10. *Let $\mathbb{S}$ be a system and let $Y \subseteq \dim(\mathbb{S})$, then*

- *$\boxminus_Y(\mathbb{S})$ is SE*

- *$\boxplus_Y(\mathbb{S})$ is SE*

This theorem is easily follows from Theorem 4.9 and the following simple fact. Let Y, Y' be disjoint sets. and let α be one of the symbols " $\boxminus$ ", " $\boxplus$ ". Then:

$$\alpha_Y \circ \alpha'_Y = \alpha_{Y \cup Y'}$$

4.4 Down-Shifting

The down shifting operator is often used in the context of shattering. We prefer to present this operator in the terminology of oblivious sorting algorithms (See [10]). To this end, we present a system $\mathbb{S}$ by its characteristic function and assume that every $v \in C(\mathbb{S})$ has a bit b_v that is either 1 or 0 according to whether $v \in S(\mathbb{S})$ or $v \notin S(\mathbb{S})$. Our "algorithms" transform the system $\mathbb{S}$ by changing these bits.

Note that a 1-dimensional cube contains two vertices. Therefore we refer to such cubes as *edges* and to an x-cube as an *x-edge*.

For a set X, we let $\preceq$ denote the partial order on $\{0,1\}^X$ that is the product of the order $0 < 1$.

Definition 4.11. *We say that an edge e is sorted under a system $\mathbb{S}$ if:*

$$e = \{u',u''\}, \ u' \preceq u'' \ and \ b_{u'} \geq b_{u''}.$$

Note that the order is reversed. This is not a typo; rather, it follows the standard definition of down shifting.

The most basic operation of our algorithms *sorts* an edge $e = \{u,v\}$ by permuting, if needed, the contents of b_u, b_v. Note that this operation mimics a "comparator" (see [10]) of a sorting network.

Definition 4.12 (Down-shifting on x)**.** *Let $\mathbb{S}$ be a system and let $x \in \dim(\mathbb{S})$.* **Down shifting** $\mathbb{S}$ **on** *x is the operation of sorting every x-edge. The resulting system is denoted $\dn_x(\mathbb{S})$.*

Lemma 4.13. *Let $\mathbb{S}$ be a system and let $x \in \dim(\mathbb{S})$. Let $\{\mathbb{D}', \mathbb{D}''\}$ be the pair of derivatives of $\mathbb{S}$ associated with x and let $\{\mathbb{R}', \mathbb{R}''\}$ be the pair of restrictions of $\dn_x(\mathbb{S})$ associated with x. Then:*

$$\{\mathbb{R}', \mathbb{R}''\} = \{\mathbb{D}', \mathbb{D}''\}$$

Definition 4.14. *Let $\mathbb{S}$ be a system.*

- *We say that $\mathbb{S}$ is x-**sorted** if every x-edge is sorted under $\mathbb{S}$.*

- *We say that $\mathbb{S}$ is **edge-sorted** if every edge is sorted under $\mathbb{S}$.*

The following lemma is easy:

Lemma 4.15. *Let $\mathbb{S}$ be a system and let $x \in \dim(\mathbb{S})$ such that $\mathbb{S}$ is x-sorted and let $\{\mathbb{S}', \mathbb{S}''\}$ be the pair of restrictions of $\mathbb{S}$ associated with x. Then:*

1. $\str(\mathbb{S}') \oplus_x \str(\mathbb{S}'') = \str(\mathbb{S})$

2. $\sstr(\mathbb{S}') \oplus_x \sstr(\mathbb{S}'') = \sstr(\mathbb{S})$

It is sometimes desired to translate our systems to the "classical" format of set-system. To this end:

Definition 4.16. *Let $\mathbb{S}$ be a system.*

$$\mathrm{SET}(\mathbb{S}) \overset{def}{=} \{f^{-1}(1) \mid f \in S(\mathbb{S})\}$$

Observe that a system $\mathbb{S}$ is edge-sorted if and only if $\mathrm{SET}(\mathbb{S})$ is closed under the subset relation. Using this observation is not hard to prove that:

Lemma 4.17. *Let $\mathbb{S}$ be an edge-sorted system. Then:*

$$\sstr(\mathbb{S}) = \str(\mathbb{S}) = \mathrm{SET}(\mathbb{S})$$

"Edge-sorting" a system by applying a sequence of down-shifts is used often in the literature. The theory of sorting teaches us how many shiftings are required to make the system edge-sorted, as indicated in the next known lemma (See [10, 7]):

Lemma 4.18. *Let $A \in K^{m \times n}$ be a 2-dimensional array with keys from a linear ordered set K. If the rows of A are sorted, then sorting the columns of A preserves the rows sorted.*

This fact is easily generalized to multi-dimensional arrays. In our context, it is manifested by the following lemma.

Lemma 4.19. *Let $\mathbb{S}$ be a system and let $x, y \in \dim(\mathbb{S})$, then $\mathrm{dn}_x(\mathrm{dn}_y(\mathbb{S}))$ is y-sorted.*

Definition 4.20. *Let Q be a sequence of operators and let x be an object. Let $Q(x)$ denote the result of applying the operators in Q on x one after the other.*

Definition 4.21. *The phrase "Q is a sequence of W" means that W is a set and Q is a sequence of members of W in which each member appears exactly once.*

An immediate corollary of Lemma 4.19 is the following fact which was observed by several authors:

Theorem 4.22. *Let $\mathbb{S}$ be a system and let Q be a sequence of $\{\mathrm{dn}_x : x \in \dim(\mathbb{S})\}$. Then $Q(\mathbb{S})$ is edge-sorted.*

4.4.1 Shifting preserves *SE*

The following theorem was proven by several authors:

Theorem 4.23. *Let $\mathbb{S}$ be a system such that $SE(\mathbb{S})$ and let $x \in \dim(\mathbb{S})$. Then*

$$\mathrm{dn}_x(\mathbb{S}) \text{ is SE}$$

Proof. Let $\mathbb{R}'$ and $\mathbb{R}''$ denote the two restricitons of $\mathrm{dn}_x(\mathbb{S})$ associated with x. By Theorem 3.14, it is enough to prove that

1. $SE(\mathbb{R}')$ and $SE(\mathbb{R}'')$,

2. $\mathrm{str}(\mathbb{R}') \oplus_x \mathrm{str}(\mathbb{R}'') = \mathrm{str}(\mathrm{dn}_x(\mathbb{S}))$.

Statement 1 easily follows from Lemma 4.13 and Theorem 4.9. Statement 2 follows from Lemma 4.15   □

5   Commutativity between Boolean operators

Henceforth, "Boolean operators" means operators of the form $\biguplus_Y$ and $\bigcap_Y$. This chapter studies the commutativity between Boolean operators and its relations to the *SE* property.

Since these operators are partial - it is important to note the following:

- Let $f : A \to B, g : A' \to B'$ be two operators. The operator $f \circ g$ is **always** defined.[7]

- It is convenient that statements of the form "$e_1 = e_2$" are always meaningful, even when e_1 or e_2 are undefined. To this end the statement "$e_1 = e_2$" is considered to be true either when both sides are defined and equal or when both are undefined. In any other case the statement is considered to be false.

The following lemma summerizes the trivial cases regarding commutativity of the Boolean operators.

Lemma 5.1. *Let Y, Y' be sets and let each of α, β be one of the symbols "$\bigcap$", "$\biguplus$".*

1. *If $Y \cap Y' \neq \emptyset$ then $\alpha_Y \circ \beta_{Y'} = \beta_{Y'} \circ \alpha_Y$ (both of them are nowhere defined).*

2. *$\alpha_Y \circ \alpha_{Y'} = \alpha_{Y'} \circ \alpha_Y$. Moreover: if $Y \cap Y' = \emptyset$ then $\alpha_Y \circ \alpha_{Y'} = \alpha_{Y \cup Y'}$*

   It is left to study the case of $Y \cap Y' = \emptyset$.

5.1   Boolean operators commute on *SE* systems

Welzl et al have shown in [8] the following theorem:

Theorem 5.2. *Let $\mathbb{S}$ be a maximum system and let $x, y \in \dim(\mathbb{S})$ such that $x \neq y$. Then:*

$$\underset{\{x\}}{\bigcap} \circ \underset{\{y\}}{\biguplus}(\mathbb{S}) = \underset{\{y\}}{\biguplus} \circ \underset{\{x\}}{\bigcap}(\mathbb{S})$$

We show the following stronger lemma:

Lemma 5.3. *Let $\mathbb{S}$ be a system and let $x, y \in \dim(\mathbb{S})$ then:*

1. *$\bigcap_{\{x\}} \circ \biguplus_{\{y\}}(\mathbb{S}) \supseteq \biguplus_{\{y\}} \circ \bigcap_{\{x\}}(\mathbb{S})$*

2. *$SE(\mathbb{S}) \implies \bigcap_{\{x\}} \circ \biguplus_{\{y\}}(\mathbb{S}) = \biguplus_{\{y\}} \circ \bigcap_{\{x\}}(\mathbb{S})$*

   Remark: Statement 1 was established by Welzl et al in [8].

Proof. We consider two cases.

Case 1: $\dim(\mathbb{S}) = \{x, y\}$. There are exactly 16 systems under this case and it is easy to verify that the lemma holds for all of them.

Case 2: $|\dim(\mathbb{S})| > 2$. Let $f = \bigcap_{\{x\}} \circ \biguplus_{\{y\}}, g = \biguplus_{\{y\}} \circ \bigcap_{\{x\}}$ and let $v \in \{0, 1\}^{\dim(\mathbb{S}) - \{x,y\}}$. It is easy to see that there exist an $\{x, y\}$-cube $C \subseteq C(\mathbb{S})$ such that:

$$f(\mathbb{S} \restriction_C) = f(\mathbb{S}) \restriction_{\{v\}},$$

---

[7]If $Range(g) \cap Dom(f) = \emptyset$ then $f \circ g$ is the operator which is nowhere defined (i.e $Dom(f \circ g) = \emptyset$)

$$g(\mathbb{S}{\restriction}_C) = g(\mathbb{S}){\restriction}_{\{v\}}$$

Thus, the lemma follows by Case 1 and Theorem 4.4. $\square$

The following lemma is an easy corollary of the first conclusion of Lemma 5.3:

Lemma 5.4. *Let Q be a sequence of operators from $\{ \bigcap_{\{x\}}, \biguplus_{\{x\}} \mid x \in \dim(\mathbb{S}) \}$ such that $Q(\mathbb{S})$ is defined. Let $\overline{Q}$ be a permutation of Q in which all the unions appear at the beginning and let $\underline{Q}$ be a permutation of Q in which all the intersections appear at the beginning. Then:*

$$\underline{Q}(\mathbb{S}) \subseteq Q(\mathbb{S}) \subseteq \overline{Q}(\mathbb{S})$$

Definition 5.5. *Let Q be a sequence of operators and let x be an object. We say that Q commutes on x if for every Q', a permutation of Q:*

$$Q(x) = Q'(x)$$

From 4.10, 5.1 and the second conclusion of Lemma 5.3 we get:

Lemma 5.6. *If $\mathbb{S}$ is an SE system and Q is a sequence of Boolean operators, then Q commutes on $\mathbb{S}$.*

Lemma 5.3 by itself does not immediately imply 5.6. Namely, having $\bigcap_{\{x\}}$ and $\biguplus_{\{y\}}$ commute on $\mathbb{S}$ for each pair $x, y \in \dim(\mathbb{S})$ does not imply that every sequence Q of such operators commutes on $\mathbb{S}$. (See Theorem 7.11)

5.2 A characterization of *Shattering* and *Strong-Shattering*

To complete the picture, what is left is to study the other direction: when does commutativity of Boolean operators on $\mathbb{S}$ imply $SE(\mathbb{S})$. To this end, we present the following characterization of shattering and strong-shattering:

Theorem 5.7. *Let $\mathbb{S}$ be a system and let $Y \subseteq X = \dim(\mathbb{S})$. Then:*

1. *$\mathbb{S}$ shatters $Y \iff \bigcap_Y \circ \biguplus_{X-Y}(\mathbb{S}) = \mathcal{K}_1$*

2. *$\mathbb{S}$ strongly shatters $Y \iff \biguplus_{X-Y} \circ \bigcap_Y(\mathbb{S}) = \mathcal{K}_1$*

Recall that $\mathcal{K}_1$ is the non-empty $\emptyset$-system
Theorem 5.7 follows from the next two straightforward lemmas.

Lemma 5.8. *For every system $\mathbb{S}$, the following statements hold:*

1. *$\biguplus_{\dim(\mathbb{S})}(\mathbb{S}) = \mathcal{K}_1 \iff S(\mathbb{S}) \neq \emptyset$*

2. *$\bigcap_{\dim(\mathbb{S})}(\mathbb{S}) = \mathcal{K}_1 \iff S(\mathbb{S}) = C(\mathbb{S})$*

Lemma 5.9. *For every system $\mathbb{S}$ and $Y \subseteq \dim(\mathbb{S})$:*

1. *$Y \in \mathrm{str}(\mathbb{S}) \iff S(\biguplus_{\dim(\mathbb{S})-Y}(\mathbb{S})) = C(\biguplus_{\dim(\mathbb{S})-Y}(\mathbb{S}))$*

2. *$Y \in \mathrm{sstr}(\mathbb{S}) \iff S(\bigcap_Y(\mathbb{S})) \neq \emptyset$*

5.3   A characterization of *SE*

Our characterization in Theorem 5.7 combined with Lemma 5.6 gives a new and simple proof of the following theorem.

Theorem 5.10 ([2])**.** *The following statements are equivalent for any system* $\mathbb{S}$*:*

1. *$SE(\mathbb{S})$*

2. *$\forall Y', Y'' \subseteq \dim(\mathbb{S})$: $\boxplus_{Y'}, \boxplus_{Y''}$ commute on $\mathbb{S}$*

6   Commutativity between Shifting operators

6.1   Shifting operators commute on *SE* systems

Lemma 6.1. *Let $\mathbb{S}$ be a SE system and let $x, y \in \dim(\mathbb{S})$ such that $x \neq y$. Then dn_x and dn_x commute on $\mathbb{S}$.*

Proof. We consider two cases.

Case 1: $\dim(\mathbb{S}) = \{x, y\}$. There are exactly 16 systems under this case and it is easy to verify that the lemma holds for all of them.

Case 2: $|\dim(\mathbb{S})| > 2$. let $f = \mathrm{dn}_x \circ \mathrm{dn}_y$, $g' = \mathrm{dn}_y \circ \mathrm{dn}_x$. It is enough to prove that for every $\{x, y\}$-cube, $C \subseteq C(\mathbb{S})$:

$$f(\mathbb{S})\!\restriction_C = g(\mathbb{S})\!\restriction_C$$

It is easy to see that for every such C:

$$f(\mathbb{S})\!\restriction_C = f(\mathbb{S}\!\restriction_C),$$

$$g(\mathbb{S})\!\restriction_C = g(\mathbb{S}\!\restriction_C)$$

Thus, the desired equality follows by Case 1 and Theorem 4.4. $\square$

From 4.23 and Lemma 6.1 we get:

Lemma 6.2. *If $\mathbb{S}$ is an SE system, and Q is a sequence of operators from $\{\mathrm{dn}_x \mid x \in \dim(\mathbb{S})\}$ then Q commutes on $\mathbb{S}$.*

6.2   A characterization of *Shattering* and *Strong-Shattering*

In this section we present a characterization of str and sstr in terms of shiftings. Recall from Definition 4.21 that he phrase "Q is a sequence of W" means that W is a set and Q is a sequence of members of W in which each member appears exactly once.

Definition 6.3. *For a set X, define:*

$$\mathrm{fullshift}(X) \stackrel{def}{=} \{Q \; : \; Q \text{ is a sequence of } \{dshift_x \; : \; x \in X\}\}$$

The following theorem is the main theorem of this section.

Theorem 6.4. *For every system $\mathbb{S}$:*

   1.
$$\mathrm{str}(\mathbb{S}) = \mathrm{SET}(\;\biguplus\; \{Q(\mathbb{S}) \; : \; Q \in \mathrm{fullshift}(\dim(\mathbb{S}))\})$$

   2.
$$\mathrm{sstr}(\mathbb{S}) = \mathrm{SET}(\;\biguplus\; \{Q(\mathbb{S}) \; : \; Q \in \mathrm{fullshift}(\dim(\mathbb{S}))\})$$

The following definition will be useful:

Definition 6.5. *Let x be an object and $i \in \{0,1\}$. Define:*

$$\beta(x,i) \stackrel{def}{=} \begin{cases} \uplus_{\{x\}} & \text{if } i = 0 \\ \sqcap_{\{x\}} & \text{if } i = 1 \end{cases}$$

Definition 6.6. *Let $\mathbb{S}$ be a system and let $x \in \dim(\mathbb{S})$.*
Let $C_i \stackrel{def}{=} \{v \in C(\mathbb{S}) \;:\; v(x) = i\}$, where $i \in \{0,1\}$ denote the two sub-cubes of $C(\mathbb{S})$ associated with x.
Define:

$$\mathbb{S}\!\restriction_{x=i} \stackrel{def}{=} \mathbb{S}\!\restriction_{C_i}, \;\; i \in \{0,1\}$$

The following lemma is straightforward.

Lemma 6.7. *Let x, y be distinct elements and let $i, j \in \{0,1\}$. Then:*

1. *The two operators $\restriction_{x=i}$ and $\restriction_{y=j}$ commute.*

2. *The two operators $\restriction_{x=i}$ and dn_y commute*

3. *$\restriction_{x=i} \circ \mathrm{dn}_x = \beta(x,i)$*

Definition 6.8. *Let X be a set, let $Q \in \mathrm{fullshift}(X)$ and let $v \in \{0,1\}^X$. Define $Q_{[v]}$ to be the sequence of Boolean operators that is derived from Q by replacing every dn_x by $\beta(x, v(x))$.*

Lemma 6.9. *Let $\mathbb{S}$ be a system, let $v \in C(\mathbb{S})$ and let $Q \in \mathrm{fullshift}(\dim(\mathbb{S}))$. Then:*

$$Q(\mathbb{S})\!\restriction_{\{v\}} = Q_{[v]}(\mathbb{S})$$

Proof. Let $X = \dim(\mathbb{S})$ and let Z be a sequence of $\{\restriction_{x=v(x)} \;:\; x \in X\}$. It is easy to see that for every X-system $\mathbb{G}$:

$$Z(\mathbb{G}) = \mathbb{G}\!\restriction_{\{v\}}.$$

Thus, $Z(Q(\mathbb{S})) = Q(\mathbb{S})\!\restriction_{\{v\}}$. Also, by Lemma 6.7:

$$Z \circ Q = Q_{[v]}.$$

Therefore, $Z(Q(\mathbb{S})) = Q_{[v]}(\mathbb{S})$. This implies the desired equality. $\qquad\square$

The following two lemmas finish the proof of Theorem 6.4:

Lemma 6.10. *For every system $\mathbb{S}$:*

$$\mathrm{str}(\mathbb{S}) = \mathrm{SET}(\; \uplus \{Q(\mathbb{S}) \;:\; Q \in \mathrm{fullshift}(\dim(\mathbb{S}))\})$$

Proof. Let $v \in C(\mathbb{S})$. It is sufiicient to prove that:

$$v^{-1}(1) \in \mathrm{str}(\mathbb{S}) \iff \uplus \{Q(\mathbb{S}) \;:\; Q \in \mathrm{fullshift}(\dim(\mathbb{S}))\}\!\restriction_{\{v\}} = \mathbb{K}_1.$$

Indeed:

$$\biguplus \{Q(\mathbb{S}) \; : \; Q \in \text{fullshift}(\dim(\mathbb{S}))\}\!\restriction_{\{v\}} =$$

by the definition of $\biguplus$

$$\biguplus \{Q(\mathbb{S})\!\restriction_{\{v\}} \; : \; Q \in \text{fullshift}(\dim(\mathbb{S}))\} =$$

by Lemma 6.9

$$\biguplus \{Q_{[v]}(\mathbb{S}) \; : \; Q \in \text{fullshift}(\dim(\mathbb{S}))\} =$$

by Lemma 5.4

$$\bigoplus_{v^{-1}(1)} (\biguplus_{v^{-1}(0)} (\mathbb{S})).$$

By Theorem 5.7 the lemma follows.   □

Using similar arguments we also prove that:

Lemma 6.11. *For every system* $\mathbb{S}$*:*

$$\text{str}(\mathbb{S}) = \text{SET}(\bigoplus \{Q(\mathbb{S}) \; : \; Q \in \text{fullshift}(\dim(\mathbb{S}))\})$$

6.3   A characterization of *SE*

Theorem 4.23 and Theorem 6.4 imply the following theorem:

Theorem 6.12 ([3, 2])**.** *The following two properties of a system* $\mathbb{S}$ *are equivalent:*

1. $\mathbb{S}$ *is SE*

2. *For every* $Q', Q'' \in \text{fullshift}(\mathbb{S})$*:*

$$Q'(\mathbb{S}) = Q''(\mathbb{S})$$

7   A weaker form of Shattering Extremality

This chapter introduces a weaker form of Shattering Extremality which is denoted by *k-Shattering Extrmality* (or *k-SE*) where k is a natural number. We study this property via the concept of *local-operators* which is interesting in its own right. The first section introduces the concept of local operators and discusses some basic properties of these operators. The second section defines *k-SE* and proves generalizations of Theorems 6.12 and 5.10 regarding equivalence of *k-SE* with weaker forms of commutativity of Boolean operators and of down-shiftings operators.

7.1   Local operators

Let *SYS* denote the set of all systems. For a system $\mathbb{S}$ and $Y \subseteq \dim(\mathbb{S})$, the collection of Y-cubes of $C(\mathbb{S})$ forms a partition of $C(\mathbb{S})$. This partition induces a partition of $\mathbb{S}$ in which every component corresponds to a restriction of $\mathbb{S}$ to a Y-cube. Informally, a Y-*local operator* is a mapping from *SYS* to *SYS* whose behaviour on $\mathbb{S}$ depends only on its behaviour on the restrictions of $\mathbb{S}$ to Y-cubes. Surprisingly, all the operators discussed in the previous chapters ($\uplus_Y, \Cap_Y, \dn_x, \restriction_{x=i}$ and $\neg$) are "highly-local" (this will become formal later).

Let X be a set. It is convenient here to extend the definition of "C is a Y-cube of $\{0,1\}^X$" to the case where $Y \not\subseteq X$. So, a new Y *cube of* $\{0,1\}^X$ when $Y \not\subseteq X$ is an $X \cap Y$-cube of $\{0,1\}^X$ under the original terminology. It is also useful to extend the definition of the operators $\restriction_{x=0}$ and $\restriction_{x=1}$ in the following way: For a system $\mathbb{S}$ such that $x \notin \dim(\mathbb{S})$, define $\mathbb{S}\restriction_{x=0} = \mathbb{S}\restriction_{x=1} = \langle \emptyset, \{0,1\}^{\dim(\mathbb{S})} \rangle$. This way, these operators are total operators from *SYS* to *SYS*.

Definition 7.1. *Let Y be a set. An operator $\alpha : SYS \to SYS$ is called Y-local if the following statements hold:*

1. *α is total.*

2. *There exists a set $Z \subseteq Y$ such that for every system $\mathbb{S}$: $\dim(\alpha(\mathbb{S})) = \dim(\mathbb{S}) - Z$.*

3. *α commutes with the operators $\restriction_{x=0}$ and $\restriction_{x=1}$ for all $x \notin Y$.*

The reader might wonder how does this definition captures the concept of locality. To answer this question, assume, for simplicity, that $Y \subseteq \dim(\mathbb{S})$ and that $Z = \emptyset$ (namely, for all $\mathbb{S}$: $\dim(\alpha(\mathbb{S})) = \dim(\mathbb{S})$). Note that every restriction of $\mathbb{S}$ to a Y-cube can be described by an appropriate sequence of operators of $\{\restriction_{x=i} : x \in \dim(\mathbb{S}) - Y\}$. Thus, item 3 in the definition of Y-local amounts to:

$$\text{For every } C, \text{ a } Y\text{-cube of } C(\mathbb{S}): \alpha(\mathbb{S})\restriction_C = \alpha(\mathbb{S}\restriction_C).$$

Therefore, each component in the partitioning of $\alpha(\mathbb{S})$ to the Y-cubes of $C(\mathbb{S})$ depends only on the corresponding component of $\mathbb{S}$. Thus, the local behaviour of α on Y-systems determines its behaviour on every system. The fact that every Y-local operator is determined by its behaviour on Y-systems is expressed in the following straightforward lemma.

Lemma 7.2. *Let* α *and* β *be two Y-local operators and let* $\mathbb{S}$ *be a system. Then the following statements are equivalent:*

1. $\alpha(\mathbb{S}) = \beta(\mathbb{S})$

2. *For every Y-cube of* $C(\mathbb{S})$, C: $\alpha(\mathbb{S}\!\restriction_C) = \beta(\mathbb{S}\!\restriction_C)$

Example 7.3. *The operator* $\neg$ *and the Identity operator are* $\emptyset$*-local with* $Z = \emptyset$.

Example 7.4. *Let x be an object. The operators* $\restriction_{x=0}$ *and* $\restriction_{x=1}$ *are* $\{x\}$*-local with* $Z = \{x\}$.

Example 7.5. *The operator* dn_x *is not total since it is not defined on systems* $\mathbb{S}$ *with* $x \notin \dim(\mathbb{S})$. *Thus, henceforth this operator is extended to such systems by* $\mathrm{dn}_x(\mathbb{S}) = \langle\emptyset, \{0,1\}^{\dim(\mathbb{S})}\rangle$. *Under this modification,* dn_x *is* $\{x\}$*-local with* $Z = \emptyset$.

Example 7.6. *The operator* $\alpha_{\{x\}}$ *where* α *is one of the symbols '* $\uplus$ *' or '* $\cap$ *', is not defined on systems* $\mathbb{S}$ *for which* $x \notin \dim(\mathbb{S})$. *Thus, henceforth this operator is extended to such systems by* $\alpha_{\{x\}}(\mathbb{S}) = \langle\emptyset, \{0,1\}^{\dim(\mathbb{S})}\rangle$. *Under this modification,* $\alpha_{\{x\}}$ *is* $\{x\}$*-local with* $Z = \{x\}$.

The following lemma is straightforward.

Lemma 7.7. *Let X and Y be sets and let* α *be an X-local operator and* β *be a Y-local operator. Then* $\alpha \circ \beta$ *is* $X \cup Y$*-local.*

Using Lemma 7.7 we can compute the locality of sequences of local operators.

Let Q be a sequence of Boolean operators. We say that Q is *meaningful* if every two distinct operators, q', q'', in Q satisfy[8] $\dim(q') \cap \dim(q'') = \emptyset$. Note that if Q is a sequence of Boolean opearators which is not meaningful then the composition of the elements of Q is the operator which is nowhere defined. Thus, all the theorems and lemmas we have proved regarding sequences of Boolean operators remain true if we consider only meaningful sequences. For a sequence of Boolean operator Q, define

$$\dim(Q) \overset{def}{=} \bigcup_{q \text{ is in } Q} \dim(q)$$

Thus, by Lemma 7.7 we obtain:

Example 7.8. *Let Q be a meaningful sequence of Boolean operators. Then Q is* $\dim(Q)$*-local.*

Similarly, we have for down-shifting operators:

Example 7.9. *Let X be a set and let* $Q \in \mathrm{fullshift}(X)$. *Then Q is X-local.*

7.2   *k-SE*

Definition 7.10. *Let* $\mathbb{S}$ *be a system and let* $k \in \mathbb{N}$. $\mathbb{S}$ *is called* $k - SE$ *if:*

For every cube of $C(\mathbb{S})$, C *with* $|\dim(C)| \leq k$: $SE(\mathbb{S}\!\restriction_C)$

We will refer to the statement "$\mathbb{S}$ *is* $k - SE$*" by "*$SE_k(\mathbb{S})$*"*

---

[8]Recall from Definition 4.6 that $\dim(\cap_Y) = \dim(\uplus_Y) = Y$.

7.2.1   k-SE and Boolean operators

The following theorem is the generalization of Theorem 5.10 concerning *k-SE*.

Theorem 7.11. *Let $\mathbb{S}$ be a system and let $k \in \mathbb{N}$. The following statements are equivalent:*

1. *Every meaningful sequence of Boolean operators Q with $|\dim(Q)| \leq k$ commute on $\mathbb{S}$.*

2. *For all disjoint Y', Y'' such that $|Y' \cup Y''| \leq k$: $\boxminus_{Y'}$ and $\uplus_{Y''}$ commute on $\mathbb{S}$.*

3. *$SE_k(\mathbb{S})$.*

Proof. $1 \iff 2$ is a simple corollary of Lemma 5.4.

$2 \iff 3$: By Lemma 7.2 and Example 7.8, Item 2 is equivalent to the following statement.

For every C, a X-cube of $C(\mathbb{S})$ where $|X| \leq k$: $\boxminus_Y$ and $\uplus_{X-Y}$ commute on $\mathbb{S}\!\restriction_C$

By Theorem 5.10, the above is equivalent to $SE_k(\mathbb{S})$. $\square$

7.2.2   *k-SE* and down-shifting operators

The following theorem generalizes Theorem 6.4.

Theorem 7.12 (Refined commution of down-shifts)**.** *Let $\mathbb{S}$ be a system and let $k \in \mathbb{N}$. The following statements are equivalent:*

1. *For all $Y \subseteq \dim(\mathbb{S})$ such that $|Y| \leq k$ and for every $Q', Q'' \in \text{fullshift}(Y)$:*

$$Q'(\mathbb{S}) = Q''(\mathbb{S})$$

2. *$SE_k(\mathbb{S})$*

Proof. The proof easily follows Lemma 7.2, Theorem 6.4 and Example 7.9 $\square$

8   Arrangements of hyperplanes in a Euclidean space

Let H be an arrangement of hyperplanes in a Euclidean space, $V = \mathbb{R}^d$. For each hyperplane $e \in H$ let one half-plane determined by e be its *positive side* and the other its *negative side*. The hyperplanes of H cut V into open regions (*cells*). Corresponding to each cell c there is $f_c \in \{+,-\}^H$ such that

$$f_c(e) = \begin{cases} + & \text{if } c \text{ is in the positive side of } e, \\ - & \text{if } c \text{ is in the negative side of } e. \end{cases}$$

In this chapter we analyze the system $\mathbb{S}_H \stackrel{def}{=} \langle \{f_c \;:\; c \text{ is a cell of } H\}, \{+,-\}^H \rangle$. Section 8.1 introduces preliminaries. Sections 8.2 and 8.3 prove the following characterizations of $\mathrm{str}(\mathbb{S}_H)$ and $\mathrm{sstr}(\mathbb{S}_H)$:

(i) $X \subseteq H$ is shattered by $\mathbb{S}_H$ if and only if the normal vectors of the hyperplanes in X are linearly independent

(ii) X is strongly-shattered by $\mathbb{S}_H$ if and only if in addition to (i) there is no hyperplane $h \in E - X$ such that $\bigcap X \subseteq h$.

Note that if H is an arrangement of lines in the plane and $\{l_1, l_2\} \subseteq H$ then the first property amounts to "l_1 and l_2 are not parallel" and the second property amounts to "The intersection point of l_1, l_2 is simple[9] in H".

Section 8.4 demonstrates a usage of the above characterizations of $\mathrm{str}(\mathbb{S}_H), \mathrm{sstr}(\mathbb{S}_h)$ for counting the number of cells of some arrangements. Section 8.5 discusses a certain generalization of such systems and presents some open questions.

Figure 1 gives an example of an arrangement of 4 lines which partition the plane into 10 regions. The positive side of each line in the arrangement is shaded.

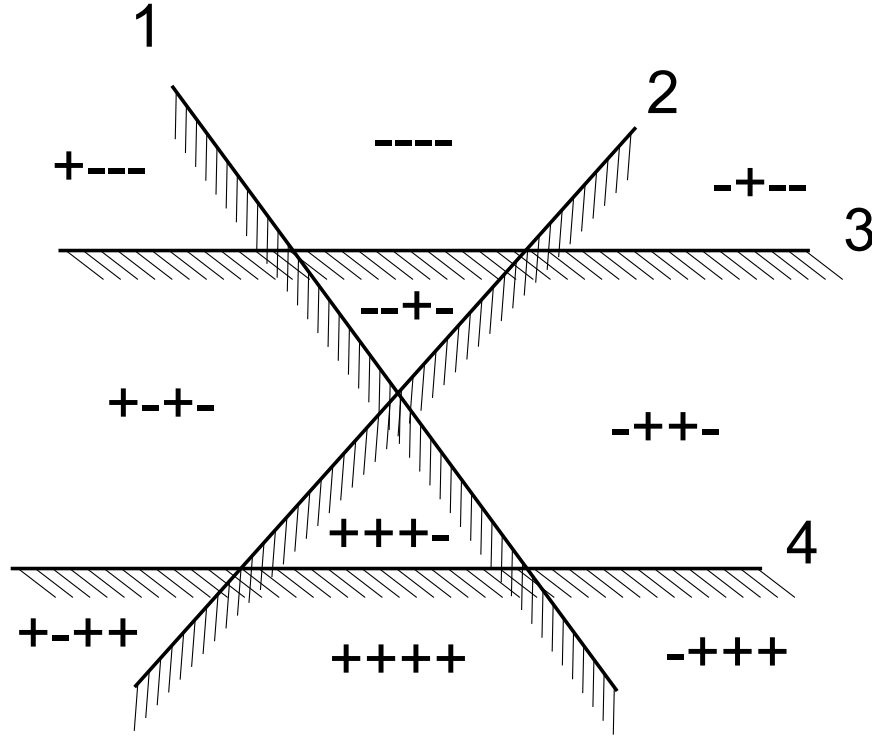[9]there is no other line in H that is incident to this point

Figure 1: An arrangement of (oriented) lines in the plane.

8.1   Notations and preliminaries

Let $V = \mathbb{R}^d$ be a Euclidean space. A *linear hyperplane*, is a $(d-1)$-dimensional sub-space of V, i.e.,

$$\{\vec{v} \in V \ : \ \vec{\alpha} \cdot \vec{v} = 0\}$$

where $\vec{\alpha}$ is a fixed nonzero vector in V and $\vec{\alpha} \cdot \vec{v}$ is the usual inner product. An *affine hyperplane*, or simply a *hyperplane*, is a translate of a linear hyperplane, i.e.

$$\{\vec{v} \in V \ : \ \vec{n} \cdot \vec{v} = r\}$$

where $\vec{n}$ is a fixed nonzero vector in V and r is a fixed scalar. $\vec{n}$ is called a *normal* of the hyperplane. Note that the normal is unique up to scaling.

An (open) *half-space* is a set $\{\vec{v} \in V \ : \ \vec{\alpha} \cdot \vec{v} > c\}$ for some $\alpha \in V$, $c \in \mathbb{R}$. If e is a hyperplane in V, then the complement $V - e$ has two components each of which is an open half-space. It is convenient to associate every (oriented) hyperplane $e \subseteq V$, with a vector $\vec{n}_e \in V$ and a number $r_e \in \mathbb{R}$ such that:

1. $e = \{\vec{x} \in V \ : \ \vec{n}_e \cdot \vec{x} = r_e\}$

2. The *positive side* of e is the set $\{\vec{x} \in V \ : \ \vec{n}_e \cdot \vec{x} > r_e\}$.

For a vector $\vec{v} \in V$, let $\text{sign}(\vec{v}) \in \{+, 0, -\}^d$ be such that, for all $i \leq d$:

$$\text{sign}(\vec{v})_i \overset{def}{=} \begin{cases} + & \text{if } v_i > 0, \\ 0 & \text{if } v_i = 0, \\ - & \text{if } v_i < 0. \end{cases}$$

For $A \subseteq V$, let $ch(A)$ denote the *convex hull* of A.

Let H be an arrangement of hyperplanes in V. We use $\bigcap H$ and $\bigcup H$ to denote the intersection and union of the hyperplanes in H.

Definition 8.1. *H is called **independent (dependent)** if the set $\{\vec{n}_e \; : \; e \in H\}$ is an independent (dependent) set of vectors.*

Lemma 8.2. *If H is independent then $\bigcap H \neq \emptyset$*

Proof. Let $\vec{x} \in V$. By the definition of $\vec{n}_e$:

$$\vec{x} \in \bigcap H \iff \forall e \in H : \vec{n}_e \cdot \vec{x} = r_e.$$

Thus it is enough to prove that there exists a solution, $\vec{x}$, for the linear system:

$$n_e \cdot x = r_e, \; \forall e \in H.$$

Independence of the n_e's implies that there exists such an x. $\qquad\square$

Definition 8.3. *Let $Y \subseteq H$. Y is called H-**regular** if*

$$Y = \{e \in H \; : \; \bigcap Y \subseteq e\}.$$

Lemma 8.4. *Let $Y \subseteq H$ such that $\bigcap Y \neq \emptyset$ and let h be a hyperplane such that $\bigcap Y \subseteq h$. Then*

$$\vec{n}_h \in \text{span}\{\vec{n}_e \; : \; e \in Y\}.$$

Proof. Let $U \overset{def}{=} \text{span}\{\vec{n}_e \; : \; e \in Y\}$. It is easy to verify that $\vec{n}_h \in (U^\perp)^\perp = U$. $\qquad\square$

Lemma 8.5. *Let $Y \subseteq E$ be independent and H-regular. Then:*

$$\bigcap Y - \bigcup (H - Y) \neq \emptyset.$$

Proof. By Lemma 8.4, there exists no $e \in H - Y$ such that $\bigcap Y \subseteq e$. Thus, every $e \in H - Y$ intersects $\bigcap Y$ in a proper affine sub-space, A_e, of $\bigcap Y$. Thus, $\{A_e \; : \; e \in H - Y\}$ is a finite collection of proper affine sub-spaces of $\bigcap Y$ and thus, by Baire Theorem, cannot cover it. $\qquad\square$

8.2   Geometric interpretation of shattering

Let H be an arrangement of hyperplanes in $\mathbb{R}^d$ and let $\mathbb{S}_H$ be the corresponding system. This section proves and discusses the following theorem:

Theorem 8.6. *Let $X \subseteq H$. Then the following statements are equivalent*

    1. $\mathbb{S}_H$ *shatters X.*

    2. *X is independent.*

Lemma 8.7. *Let $X \subseteq H$. Then*

$$X \text{ is independent} \implies \mathbb{S}_H \text{ shatters } X.$$

Proof. It is enough to show that for every $f \in \{+,-\}^X$ there exists some cell c such that f_c and f agree on X. Let $f \in \{+,-\}^X$. Define

$$y_e \stackrel{def}{=} \begin{cases} r_e + 1 & \text{if } f(e) = + \\ r_e - 1 & \text{if } f(e) = - \end{cases}$$

Since X is independent the linear system

$$\vec{n}_e \cdot \vec{x} = y_e, e \in X$$

has a solution in $\mathbb{R}^d$. Let c denote a cell whose closure, $\overline{c}$, contains such a solution. Then f_c and f agrees on X as required. □

Proving the other requires more work. We split it to two parts; first, it is shown that if $X \subseteq H$ is shattered by $\mathbb{S}_H$ then $\bigcap X \neq \emptyset$. In the second part we establish that if $\mathbb{S}_H$ shatters X then X is independent.

For the first part, the following lemma is useful.

Lemma 8.8. *Let $k \in \mathbb{N}$ and let $A \subseteq \mathbb{R}^k$ be a set that satisfies:*

$$(\forall f \in \{+,-\}^k)(\exists v \in A): \ \text{sign}(v) = f.$$

Then $\vec{0} \in ch(A)$.

Lemma 8.8 is easily proved by induction on k.

Lemma 8.9. *Let $X \subseteq H$ be a non-empty set. Then*

$$X \in \text{str}(\mathbb{S}_H) \implies \bigcap X \neq \emptyset.$$

Proof. Let $A \in \mathbb{R}^{X \times d}$ be a matrix with the $\vec{n}_e$'s in X as its rows and let $\vec{r} \in \mathbb{R}^X$ be a vector with the r_e's as its entries. It is enough to show that there exists x such that $A\vec{x} = \vec{r}$.

Let T denote the map $\vec{x} \mapsto A\vec{x} - \vec{r}$. T is an affine map. Since $\mathbb{S}_H$ shatters X, it follows that:

$$(\forall f \in \{+,-\}^X)(\exists x \in \mathbb{R}^d) : \text{sign}(A\vec{x} - \vec{r}) = f$$

Let $A \subseteq \mathbb{R}^d$ be a set that for every $f \in \{+,-\}^X$ contains $x_f \in \mathbb{R}^d$ for which $\text{sign}(A\vec{x} - \vec{r}) = f$. Therefore, according to Lemma 8.8, it follows that $\vec{0} \in ch(T(A))$. Since T is affine, we have that $ch(T(A)) = T(ch(A))$. Thus, there exists $\vec{x} \in ch(A)$ such that $T(\vec{x}) = 0$, which means that: $A\vec{x} = \vec{r}$ as required. □

Lemma 8.10. *Let $X \subseteq H$. Then*

$$\mathbb{S}_H \text{ shatters } X \implies X \text{ is independent.}$$

Proof. By Lemma 8.9 we may assume that for all $e \in X$: $r_e = 0$. (Otherwise, translate $\mathbb{R}^d$ such that the origin is in $\bigcap X$) Let $\{\alpha_e \; : \; e \in X\}$ be coefficients such that $\sum_{e \in X} \alpha_e \vec{n}_e = 0$. It is enough to show that $\forall e \in X : \alpha_e = 0$.

Let $f \in \{+, -\}^X$ such that

$$f(e) \stackrel{def}{=} \begin{cases} + & \text{if } \alpha_e \geq 0 \\ - & \text{otherwise} \end{cases}.$$

Since $\mathbb{S}_H$ shatters X there exists $x \in \mathbb{R}^d$ such that

$$\text{sign}(\vec{x} \cdot \vec{n}_e) = f(e)$$

Furhtermore, we may choose $\vec{x}$ such that $\forall e \leq X : |\vec{x} \cdot \vec{n}_e| \geq 1$.

$$\begin{aligned} 0 &= \vec{x} \cdot \sum_{e \in X} \alpha_e \vec{n}_e && \text{since } \sum_{e \in X} \alpha_e \vec{n}_e = 0 \\ &= \sum_{e \in X} \alpha_e \vec{x} \cdot \vec{n}_e \\ &\geq \sum_{\alpha_e \geq 0} \alpha_e - \sum_{\alpha_e < 0} \alpha_e && \text{by definition of } \vec{x} \end{aligned}$$

Therefore:

$$\sum_{\alpha_e \geq 0} \alpha_e - \sum_{\alpha_e < 0} \alpha_e = 0.$$

This is only possible if $\forall e \in X : \alpha_e = 0$. $\qquad\square$

8.3 Geometric interpretation of strong-shattering

Let H be an arrangement of hyperplanes in $\mathbb{R}^d$ and let $\mathbb{S}_H$ be the corresponding system. This section proves and discusses the following theorem:

Theorem 8.11. *Let $X \subseteq H$. Then the following statements are equivalent*

1. $\mathbb{S}_H$ *strongly-shatters X.*

2. *X is independent and H-regular.*

Lemma 8.12. *Let $X \subseteq H$. Then*

$$X \text{ is independent and } H\text{-regular} \implies \mathbb{S}_H \text{ strongly-shatters } X.$$

Proof. According to Lemma 8.5 there exists $\vec{p} \in \bigcap X - \bigcup(H - X)$. Therefore:

$$\forall e \in H - X : |\vec{n}_e \cdot \vec{p} - r_e| > 0.$$

Define $g \in \{+, -\}^{H-X}$ to be:

$$g(e) \stackrel{def}{=} \mathrm{sign}(\vec{n}_e \cdot \vec{p} - r_e)$$

It is enough to show that for all $f \in \{+, -\}^X$ there exists a cell c such that $f_c = g \star f$.
Let $f \in \{+, -\}^X$. Define

$$y_e \stackrel{def}{=} \begin{cases} 1 & \text{if } f(e) = +, \\ -1 & \text{if } f(e) = -. \end{cases}$$

Consider the linear system

$$\vec{n}_e \cdot \vec{x} = y_e \text{ for all } e \in X$$

Since X is independent, this system has a solution $\vec{d}$.
Pick $\varepsilon > 0$ sufficiently small such that

$$\forall e \in E - X : \mathrm{sign}(\vec{n}_e \cdot \vec{p} - r_e) = \mathrm{sign}(\vec{n}_e \cdot (\vec{p} + \varepsilon\vec{d}) - r_e).$$

Let c denote the cell that contains $\vec{p} + \varepsilon\vec{d}$. By the definitions of ε and $\vec{d}$ it follows that $f_c = g \star f$ as required. $\square$

For the other direction:

Lemma 8.13. *Let $X \subseteq H$. Then*

$$\mathbb{S}_H \text{ strongly-shatters } X \implies \mathbb{S}_H \text{ is independent and } H\text{-regular.}$$

Proof. We will prove the contra-position of this claim. If X is dependent then $\mathbb{S}_H$ does not shatter X. In particular, this means that $\mathbb{S}_H$ doesnt strongly-shatter X. Otherwise, assume X is independnet and H-irregular. Therefore, by Lemma 8.9 $\bigcap X \neq \emptyset$. Moreover, we may assume that $\vec{0} \in \bigcap X$ (by translating the space if needed).

Since X independent and H-irregular, it follows that there exists $h \in H - X$ such that $\bigcap X \subseteq h$. Thus, by Lemma 8.4, $\vec{n}_h \in \mathrm{span}\{\vec{n}_e : e \in X\}$. which means that there exist coefficients $\{\alpha_e \in \mathbb{R} : e \in X\}$ such that

$$\sum_{e \in X} \alpha_e n_e = n_h.$$

Note that $\vec{0} \in \bigcap X \subseteq h$ implies that $r_h = 0$. Assume, by way of contradiction that $\mathbb{S}_H$ strongly-shatters X. Let

$$g(e) \stackrel{def}{=} \begin{cases} + & \text{if } \alpha_e \geq 0 \\ - & \text{otherwise} \end{cases}$$

and

$$-g(e) \stackrel{def}{=} \begin{cases} + & \text{if } g(e) = - \\ - & \text{otherwise.} \end{cases}$$

Since X is strongly shattered by $\mathbb{S}_H$, there exist two open cells c', c'' such that:

(i) $f_{c'}$ agrees with g on X

(ii) $f_{c''}$ agrees with $-g$ on X

(iii) $f_{c'}$ agrees with f_c'' on $H - X$

We will derive a contradiction to item (iii) by showing that $f_{c'}(h) \neq f_{c''}(h)$.
Let $\vec{x'} \in c'$ and $\vec{x''} \in c''$.

$$
\begin{aligned}
f_{c'}(h) &= \text{sign}(\vec{n_h} \cdot \vec{x'}) \\
&= \text{sign}([\sum_{e \in X} \alpha_e \vec{n_e}] \cdot \vec{x'}) \qquad\qquad \sum_{e \in X} \alpha_e \vec{n_e} = \vec{n_h} \\
&= \text{sign}(\sum_{e \in X} \alpha_e [\vec{n_e} \cdot \vec{x'}]) \\
&= + \qquad\qquad\qquad\qquad \text{by definition of } f_{c'} : \alpha_e [\vec{n_e} \cdot \vec{x'}] \geq 0
\end{aligned}
$$

.

Similarly we have $f_{c''}(h) = -$ and we are done. $\qquad\qquad\qquad\qquad \square$

8.4 Counting cells

The following lemma is a corollary of Theorems 8.6, 8.11 and Lemma 8.4.

Lemma 8.14. *Let H be an arrangement of hyperplanes in $\mathbb{R}^d$. Then, the following statements are equivalent.*

1. *$SE(\mathbb{S}_H)$*

2. *$\forall X \subseteq H$: X is independent $\iff \bigcap X \neq \emptyset$.*

Let H be an arrangement of hyperplanes such that $SE(\mathbb{S}_H)$. By Theorem 3.10, this means that $|\mathbb{S}| = |\text{str}(\mathbb{S})|$. Thus, by Lemma 8.14 and Theorem 8.6 we get the following result:

Lemma 8.15. *Let H be an arrangement of hyperplanes in $\mathbb{R}^d$ such that*

$$\forall X \subseteq H: X \text{ is independent} \iff \bigcap X \neq \emptyset.$$

Let C denote the number of cells in the arrangement. Then:

$$C = |\{X \subseteq H : X \text{ is independent}\}|.$$

in particular, if the arrangement H is in a general position (i.e. every d hyperplanes intersect in a single point and no $d+1$ hyperplanes intersect), then it is easy to verify that H satisfies the premise of Lemma 8.15 and that

$$\{X \subseteq H : X \text{ is independent}\} = \{X \subseteq H : |X| \leq d\}.$$

This proves the well known result regarding the number of cells, C, in such an arrangement. Namely $C = \sum_{i=0}^{d} \binom{|X|}{i}$.

8.5   Convex systems and some questions

Let $\mathbb{S}$ be a system. $\mathbb{S}$ is called *Euclidean* if there exist a Euclidean space, $\mathbb{R}^d$, and an arrangement of hyperplanes H such that $\mathbb{S} = \mathbb{S}_H$.

Let H be an arrangement of hyperplanes in $\mathbb{R}^d$ and let $\mathbb{S}_H$ be the corresponding system. Let $e \in H$ and let $\mathbb{S}'_H, \mathbb{S}''_H$ be the two restrictions of $\mathbb{S}_H$ associated with x. It is easy to see that $\mathbb{S}'_H \uplus \mathbb{S}''_H$ corresponds to the arrangement $H' = H - \{e\}$ and that $\mathbb{S}'_H \uplus \mathbb{S}''_H$ corresponds to the arrangement H' (of one dimension smaller) obtained as the intersection of e with the remaining hyperplanes. Thus, Euclidean systems are closed under Boolean operators. It is not hard to prove that Euclidean systems are not closed under restrictions. However, the following generalization enables the representation of $\mathbb{S}'_H, \mathbb{S}''_H$ in a geometric setup.

Let K be an open convex set in $\mathbb{R}^d$. Let

$$\mathbb{S}_{H,K} = \langle \{ f_c \ : \ c \text{ is a cell of } H, \ c \cap K \neq \emptyset \}, \{+, -\}^H \rangle.$$

It is easy to see that choosing K to be one of the half-spaces determined by e yields a representation of $\mathbb{S}'_H$ and $\mathbb{S}''_H$ as $\mathbb{S}_{H,K}$.

This setup generalizes the setups investigated by Lawerence in [12] and by Welzl et al in [8]. The systems studied in [12] are obtained by choosing $H = \{ \{ \vec{x} \ : \ x_i = 0 \} \ : \ 1 \leq i \leq d \}$ and K to be arbitrary and the systems studied in [8] are obtained by choosing H to be any general-position arrangement and $K = \mathbb{R}^d$.

The following two theorems can be proved similarly to the way in which we proved theorems 8.11, 8.6

Theorem 8.16. *Let $X \subseteq H$. Then the following statements are equivalent*

1. *$\mathbb{S}_{H,K}$ shatters X.*

2. *X is independent.*

3. *$\bigcap X$ and K share a common point.*

Theorem 8.17. *Let $X \subseteq H$. Then the following statements are equivalent*

1. *$\mathbb{S}_{H,K}$ strongly-shatters X.*

2. *X is independent, H-regular.*

3. *$\bigcap X$ and K share a common point.*

The above characterizations of str, sstr in this geometric setup yields a big class of *SE* systems that rise from arrangements of hyperplanes in a Euclidean space. It is natural to ask whether every *SE* system can be implemented in such a way. The following is a partial answer to this question.

Theorem 8.18. *There exists an SE system $\mathbb{S}$ such that there exists no arrangement of hyperplanes, H, such that $\mathbb{S} = \mathbb{S}_H$.*

Proof. By Theorem 8.6, for every H, it follows that $\mathrm{str}(S_H)$ is a matroid. It is easy to construct an *SE*-family $\mathbb{S}$ such that $\mathrm{str}(s)$ is not a matroid. $\square$

However, the answer to the following question is not known.

Question 8.19. *Let $\mathbb{S}$ be an SE system. Is there an arrangement H and a convex set K such that $\mathbb{S} = \mathbb{S}_{H,K}$?*

9   Orientations of an undirected graph

This chapter presents a method which uses the tools we developed in order to prove certain equalities and inequalities involving undirected graphs and their orientations. Unlike most applications of these concepts, which use the "shattering" relation, this one demonstrates a usage of "strong-shattering". The method is demonstrated via two examples.

Henceforth, let $G = (V, E)$ be an arbitrary fixed undirected simple graph. We represent an orientation of E as a function $d : E \to \{+, -\}$. In order to enable this, pick a fixed orientation $\overrightarrow{E}$ of E. Thus, a function $d : E \to \{+, -\}$ define an orientation of E, relative to $\overrightarrow{E}$, in the obvious way: If $d(e) = +$ then e is oriented the same way as in $\overrightarrow{E}$. Else, if $d(e) = -$ then e is oriented in the opposite way as in $\overrightarrow{E}$. For an orientation d, let G_d denote the directed graph obtained by orienting E according to d.

For $X \subseteq E$, let $G_X \overset{def}{=} (V, X)$ denote the undirected subgraph of G with X as its edges.

9.1   Cyclic orientations

The first example proves the following inequality:

Theorem 9.1. *The number of orientations of G that yield a directed cycle is at least the number of its subgraphs that have an undirected cycle.*

Consider the system $\mathbb{S} \overset{def}{=} \langle \{d \in \{+, -\}^E \ : \ G_d \text{ has a directed cycle}\}, \{+, -\}^E \rangle$. The main step in this method is to characterize $\mathrm{sstr}(\mathbb{S})$ or $\mathrm{str}(\mathbb{S})$ (or both). In this example, the following lemma characterize $\mathrm{sstr}(\mathbb{S})$.

Lemma 9.2. *Let $X \subseteq E$. Then $X \in \mathrm{sstr}(\mathbb{S}) \iff G_{E-X}$ has a cycle*

Proof. If there exists a cycle, C, in $E - X$ then there exists an orientation of $E - X$, d, such that C is oriented to a directed cycle. Clearly, every extension of d to an orientation of E contains this directed cycle. This means that $X \in \mathrm{sstr}(\mathbb{S})$

For the other direction, assume that $E - X$ doesn't contain a cycle. To establish that $X \notin \mathrm{sstr}(\mathbb{S})$, it is enough to show that:

$$(\forall d' : (E - X) \to \{+, -\})(\exists d'' : X \to \{+, -\}) : G_{d' \star d''} \text{ has no directed cycle.}$$

Let $d' : (E - X) \to \{+, -\}$. Since $E - X$ has no cycles, it follows that the orientation of G_{E-X} according to d' is a DAG whose edges form a pre-order, P, on V. Pick (by topological sorting) a linear order L of V that extends P. Let $d'' : X \to \{+, -\}$ be the orientation of the remaining edges according[10] to L. Clearly, the resulting orientation, $d' \star d''$, is a-cyclic. $\qquad\square$

Theorem 9.1, easily follows from the Sandwich Theorem (Theorem 3.3). It is easy to see that in some graphs the inequality is strict, as observed by taking G to be a simple

---

[10]every edge is oriented towards the bigger vertex

cycle. Indeed - for such G there are two possible orientations that yield a directed cycle but only one $X \subseteq E$ such that G_X has a cycle.

Consider the system $\neg\mathbb{S}$, namely the orientations of G that yield an a-cyclic graph. By Lemma 2.9 it follows that

Lemma 9.3. *Let $X \subseteq E$. Then $X \in \text{str}(\neg\mathbb{S}) \iff G_X$ is a forest .*

This, combined with the Sandwich Theorem, gives the following inequality:

Theorem 9.4. *The number of orientations of G that yield an a-cyclic graph is at most the number of its subgraphs that are forests.*

These characterizations of $\text{str}(\neg\mathbb{S})$ and $\text{sstr}(\mathbb{S})$ also give a natural interpretation of the VC-dimension and the *Dual VC-dimension*. The dual VC-dimension of a system is obtained by applying the Duality Tranformation on the text defining the VC-dimension:

Definition 9.5 (Dual-VC-dimension)**.** *The* **Dual VC-dimension** *(Vapnik Chervonenkis dimension) of a system $\mathbb{S}$, denoted* $\text{dvc}(\mathbb{S})$*, is the cardinality of the largest*[11] *subset that is strongly-shattered by it.*

The next lemmas easily follows from the characterization of $\text{sstr}(\mathbb{S})$ and $\text{str}(\neg\mathbb{S})$.

Lemma 9.6. *Let $|V| = n$ and let $|E| = m$. Then*

1. *$\text{vc}(\neg\mathbb{S}) = k$ where k is the size of a maximum subforest of G.*

2. *$\text{dvc}(\mathbb{S}) = m - c$, where c is the size of a smallest cycle in G.*

9.2   Path preserving orientations

The second example will demonstrate a proof for the following equality:

Theorem 9.7. *Let $W \subseteq V$ and let $s \in W$. Then, the number of orientations of G in which all vertices in W are reachable from s equals to the number of subgraphs of G in which all vertices of W lie in the same connected component.*

Let $\mathbb{S}$ denote the system $\langle \{d \in \{+,-\}^E : \forall w \in W : G_d \text{ has an } s-w \text{ directed path}\}, \{+,-\}^E \rangle$ The following lemma characterizes $\text{str}(\mathbb{S})$ and $\text{sstr}(\mathbb{S})$. Moreover, it shows that $SE(\mathbb{S})$.

Lemma 9.8. *Let $X \subseteq E$. Then, the following three statements are equivalent:*

1. *$\mathbb{S}$ strongly shatters X*

2. *$\mathbb{S}$ shatters X*

3. *All vertices of W lie in the same connected componenet of G_{E-X}*

---

[11]As a special case, $\text{dvc}(\mathbb{S}) = -1$ when $S(\mathbb{S}) = \emptyset$

Proof. 1 $\implies$ 2 is trivial.

Consider 2 $\implies$ 3. To establish this implication, we prove its contra-position, namely we assume that not all vertices of W lie in the same connected component of $E - X$ and we show that there exist $d \in \{+, -\}^X$ such that no orientation in $\mathbb{S}$ agrees with d on X. Let $w \in W$ such that w and s lie in different components of $E - X$ and let S denote the connected componenet of S . Let d denote an orientation of X in which w is not reachable from s (to obtain such an orientation, orient all edges connecting S and $V - S$ towards S). Clearly, every orientation of E, that agrees with d on X has no directed path from S to $V - S$ and in particular no directed path from $s \in S$ to $w \in V - S$. Thus, every such orientation is not in $\mathbb{S}$.

Consider 3 $\implies$ 1. Assume that all vertices of W lie in the same connected componenet of G_{E-X}. Let d be an orientation of $E - X$ such that s is a root of its connected component. Note that in particular this means that every $w \in W$ is reachable from s. Clearly, in every extension of d to an orientation of E, every $w \in W$ is reachable from s. $\square$

Similarly to the first example, the Sandwich Theorem establishes Theorem 9.7.
The following lemma is an easy corollary of Theorem 9.8.

Lemma 9.9. $\mathrm{vc}(\mathbb{S}) = m - t$ where t is the size of a Steiner tree for W.

Consider the case in which $W = \{s, t\}$. In this case, $\mathrm{vc}(\mathbb{S}) = m - d$ where d is the size of a shortest $s - t$ path in G. Moreover, $\neg \mathbb{S}$ is the set of all orientations that do not contain an $s - t$ directed path, and by Lemma 2.9 we obtain:

Lemma 9.10. *The number of orientations that do not contain a directed path from s to t equals to the number of subsets of E that separates s and t. Moreover,* $\mathrm{vc}(\neg \mathbb{S})$ *is* $m - c$, where c is the size of a minimum $s - t$ cut.

9.2.1   An interesting class of SE systems

In the preceding chapters we showed that SE systems are preserved by many operations (e.g complementing, restrictions, Boolean operators,...). A natural question is to ask whether the class of SE systems is closed under union and intersection. It is not hard to design to SE systems, $\mathbb{S}'$ and $\mathbb{S}''$ that demonstrates that it is not the case. However, there are some subclasses of SE systems that are closed under these operations. One trivial example is the class of all edge-sorted systems. In this subsection we introduce another such class.

Pick some vertex $s \in V$ and let $v \in V - \{x\}$. Let $\mathbb{S}_{s,v}$ be the system of all orientations which yield a directed path from s to v. Consider the class

$$\{\mathbb{S}_{s,v} \; : \; v \in V - \{s\}\}$$

By Lemma 9.8 it follows that every system in this class is SE. Moreover, by the same lemma we get that any intersection of systems in this class is SE. It is not hard to prove, using the same method, that every union of elements in this class is SE. Therefore, this forms a non-trivial example for a class of SE systems that is closed under unions and intersections.

References

[1] Richard P. Anstee, Lajos Rónyai, and Attila Sali. Shattering news. *Graphs and Combinatorics*, 18(1):59–73, 2002.

[2] Hans-Jürgen Bandelt, Victor Chepoi, Andreas W. M. Dress, and Jack H. Koolen. Combinatorics of lopsided sets. *Eur. J. Comb.*, 27(5):669–689, 2006.

[3] Béla Bollobás and A. J. Radcliffe. Defect sauer results. *J. Comb. Theory, Ser. A*, 72(2):189–208, 1995.

[4] Béla Bollobás, A. J. Radcliffe, and Leader I. Reverse kleitman inequalities. *Proc. London Math. Soc., Ser. A*, (3) 58:153–168, 1989.

[5] A.W.M. Dress. Towards a theory of holistic clustering. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, 37 Amer. Math. Soc.:271–289, 1997.

[6] S. Floyd and Manfred K. Warmuth. Sample compression, learnability, and the vapnik-chervonenkis dimension. *Machine Learning*, 21(3):269–304, 1995.

[7] David Gale and Richard M. Karp. A phenomenon in the theory of sorting. *J. Comput. Syst. Sci.*, 6(2):103–115, April 1972.

[8] Bernd Gartner and Emo Welzl. Vapnik-chervonenkis dimension and (pseudo-)hyperplane arrangements. *Discrete and Computational Geometry (DCG)*, 12:399–432, 1994.

[9] G. Greco. Embeddings and the trace of finite sets. *Inf. Process. Lett.*, 67(4):199–203, 1998.

[10] Donald E. Knuth. *The art of computer programming, volume 3: (2nd ed.) sorting and searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998.

[11] Dima Kuzmin and Manfred K. Warmuth. Unlabeled compression schemes for maximum classes. *Journal of Machine Learning Research*, 8:2047–2081, 2007.

[12] J. Lawrence. Lopsided sets and orthant-intersection by convex sets. *Pac. J. Math.*, 104(1):155–173, 1983.

[13] A. Pajor. Sous-espaces l_1^n des espaces de banach. *Travaux en Cours. Hermann, Paris*, 1985.

[14] Benjamin I. P. Rubinstein, Peter L. Bartlett, and J. Hyam Rubinstein. Shifting: One-inclusion mistake bounds and sample compression. *J. Comput. Syst. Sci.*, 75(1):37–59, 2009.

[15] Benjamin I. P. Rubinstein and J. Hyam Rubinstein. A geometric approach to sample compression. *CoRR*, abs/0911.3633, 2009.

[16] J. Hyam Rubinstein and Benjamin I. P. Rubinstein. Geometric & topological representations of maximum classes with applications to sample compression. In *COLT*, pages 299–310, 2008.

[17] N. Sauer. On the density of families of sets. *J. Comb. Theory, Ser. A*, 13:145–147, 1972.

[18] S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pac. J. Math.*, 41:247–261, 1972.

[19] V. N. Vapnik and A. Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probability. *Theory Probab. Appl.*, 16:264–280, 1971.